Streaming Over HFC—MPEG-2 or IP or Both?

Mukta L. Kar and Bill Kostka
Cable Television Laboratories, Inc.
Majid Chelehmal
Terayon Communication Systems
Munsi Haque
Philips Semiconductor

Abstract

With the rapid progress in science and technology, more and more applications involving audio and video streams are emerging. The MPEG-2 compression standard, coupled with advances in digital modulation, has already made an impact in the broadcast industry by increasing the number of channels and viewing quality. However, consumers have limited choice—purchasing or renting CDs and DVDs from a limited stock or selecting from limited movie and music offerings.

Digital technology has advanced in a number of fronts, including the transport of data over public or private networks. High-speed data can be transported in a variety of ways—unicast, multicast, or even broadcast using various transport protocols. Digital information, such as audio, video, text, graphics, etc., differs only in the size of data. For example, digitized video is much larger than digitized audio. Advances in fiber technology have increased the capacity of both public and private networks to the point that audio and video streaming is becoming a reality. CableLabs, and other reseach institutes, have found that there are a few obstacles to overcome before broadcast-quality video streaming is possible over public networks. When those issues are resolved, electronic search engines will be able to find the desired content in archives located anywhere in the world. Consumers will be able to enjoy their content choices with a click of a button.

One major issue is the transport mechanism (protocol) for real-time, uninterrupted flow of digital audio, video, graphics, etc. MPEG-2 transport protocol and IP-based Real-time Transport Protocol (RTP) are currently the two leading protocols used for the delivery of digital content in real-time. This paper will analyze these two protocols in detail, and will present comparative studies for streaming technology. This technology is not mature; e.g., streaming in IP networks is implemented as part of transmission application layer protocols, where the unreliable user datagram protocol (UDP) is used mostly at the transport layer. To alleviate packet loss, RTP with quality-of-service (QOS) routing is considered for improved services in IP networks.

Issues and obstacles for streaming audio-visual content, particularly over the public networks, will be addressed in this paper. Research activities addressing some of the problems also will be discussed. Finally, content streaming based on MPEG-4, the recently completed multimedia standard, also will be discussed.

Introduction

Until a few years ago, cable systems were one-way networks used to deliver premium analog television content. Equipment used in cable plants was proprietary in nature. As a result, the majority of the equipment was not interoperable and, therefore, not portable across cable systems. Advances in digital technology have brought about a revolution in the computer and communications industries, including cable networks. Digital television signal compression technology and MPEG-2 standardization, coupled with digital modulation, have ushered in a new era in television broadcasting. One of the most important benefits is the bandwidth efficiency in spectrum utilization compared to analog broadcasting. An existing 6-MHz analog channel can be used to send multiple digital channels with equal or better picture and sound quality. This indicates that the existing limited number of analog channels can be transformed into a larger number of viewing or logical channels. These logical channels can be used for the delivery of audio, video, and high-speed data services. Digital technology has transformed the analog cable network into a broadband multimedia delivery system.

Prior to the addition of digital technology to existing analog systems, the cable industry had researched transport protocols and digital modulation systems for transmission of compressed digital audio/video over the cable networks. About the same time, the MPEG-2 standard [1] was completed by the moving picture expert group (MPEG) of the International Organization for Standardization (ISO). Unlike MPEG-1 Systems [2], which deal with a single program, MPEG-2 systems can handle multiple programs and have an added transport layer optimized to broadcast digital audio and video synchronously. Because cable systems primarily deliver premium television content, the MPEG-2 transport system became cable's primary choice. Also, hybrid fiber/coaxial (HFC) cable networks are much less susceptible to atmospheric noise; a better signal-to-noise ratio is available compared to other broadcasting systems. To take advantage of this, the cable industry chose to use a higher order digital modulation system, 64-QAM or 256-QAM (quadrature amplitude modulation), which transforms a 6-MHz physical spectrum into a larger digital bandwidth (27 Mbps for 64 QAM and 38 Mbps for 256 QAM) described in ITU standard J.83-B [3]. Digital technology vastly increases the channel capacity of the existing cable plant. This increased capacity will allow the delivery of an increased number of TV channels and other digital services as well.

In adding digital capability, cable networks are equipped for MPEG-2 transport at the baseband. Figure 1 shows a simplified diagram for delivery of digital TV signals using baseband MPEG-2 transport. The multiplexer's output is an MPEG-2 systems-compliant transport stream consisting of MPEG-2 transport packets. At the modulator, the transport packet payload, excluding the packets carrying system information, may be encrypted (optional). Forward error correction (FEC) is applied to protect against noise in the transmission channel. The resulting bitstream then digitally modulates a 6-MHz carrier. Addition of FEC and digital modulation is described in ITU Standard J.83-B. The modulated carrier is upconverted to a desired cable channel before combining with other channels for transmission over the cable network. In the upstream (return channel), baseband transport is similar to downstream transport, except that lower order modulation (e.g., QPSK) is used. Upstream transport is used for management messages and user interactivity, and is shared by many set-top boxes. QPSK provides more robust modulation so that packet loss due to noise will be minimal.

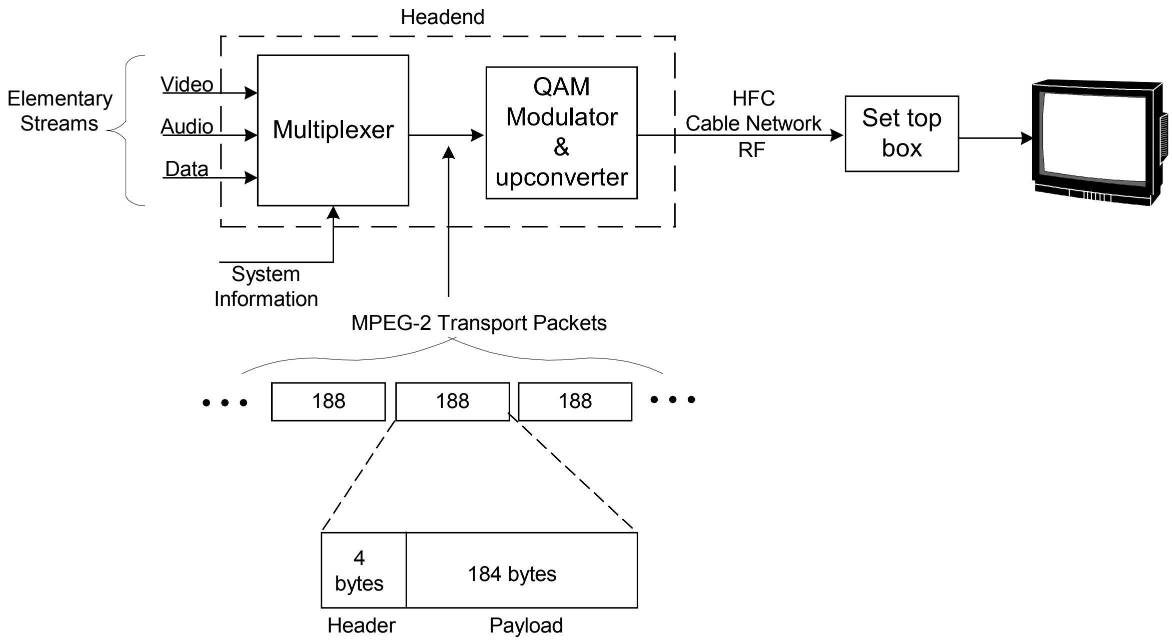


Figure 1. Broadcast of MPEG-2 audio/video over HFC cable network (downstream)

In recent years, high-speed Internet connectivity has become a popular service. Internet connectivity allows use of digital services available over the Internet, such as web-browsing, web-casting, email, home shopping, chat, etc. The cable industry decided to add Internet connectivity to cable networks, but services over the Internet require two-way networks. In adding digital technology, many cable systems have been upgraded to provide two-way connectivity. The Internet is primarily a high-speed data network and uses Internet protocol (IP) transport [11], not MPEG-2, as the baseband transport. Since the early days of cable, video delivery has been a primary source of revenue and, even today, the major part of cable systems' revenue comes from video delivery. Figure 2 shows a simple diagram of the interface between a HFC network and the Internet. CableLabs, in collaboration with its members and the vendor community, have developed an interface standard known as Data Over Cable Service Interface Specification (DOCSIS) [12]. The main objective of the DOCSIS specification is to enable many vendors to design interface equipment in a competitive market place. Equipment designed based on a standard will be interoperable, portable, and available in retail markets. Per the DOCSIS standard, IP packets are encapsulated into MPEG-2 transport packets before transmission over the network in the downstream direction. If the payload size is larger than 184 bytes, the packet is broken into smaller ones before being sent as MPEG-2 transport packet payloads. Conversion of IP packets to MPEG-2 packets is performed per following rule.

$$\text{If} \quad \mathrm{mod}(L, 184) = 0;\ N = L/184$$
$$\mathrm{mod}(L, 184) \neq 0;\ N = (L/184) + 1$$

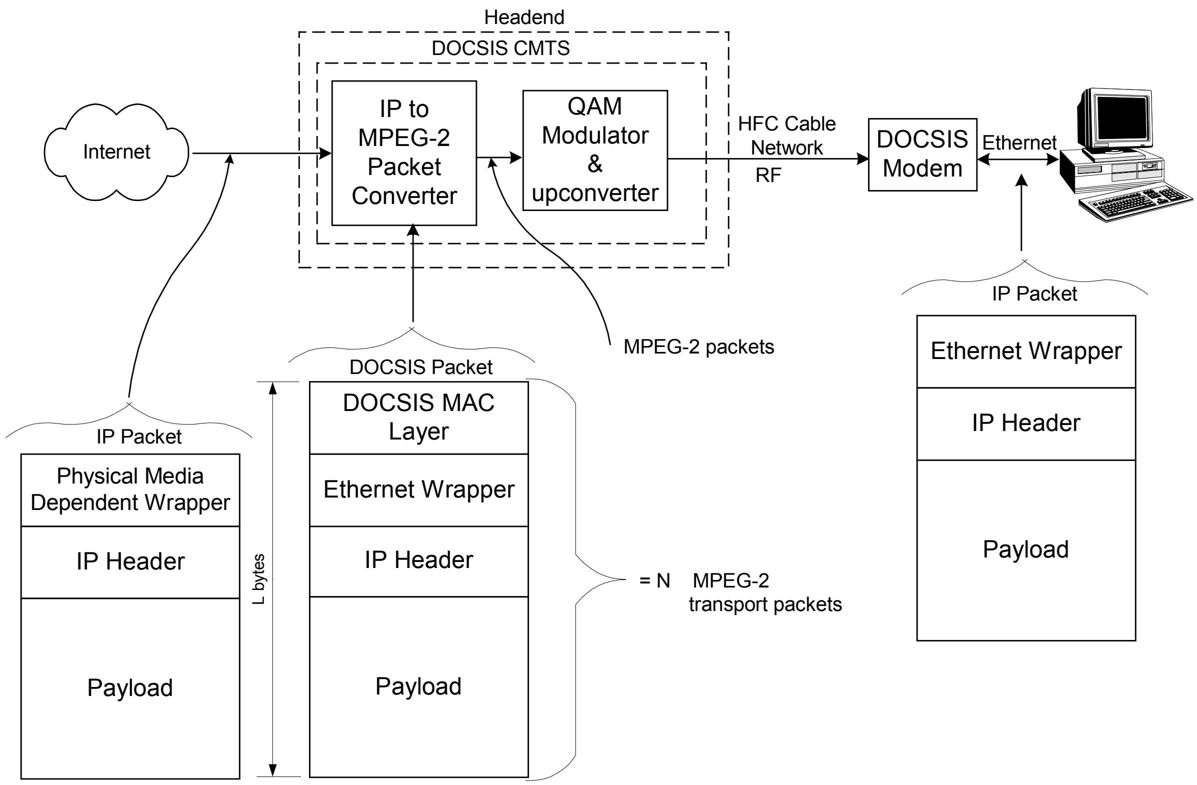


Figure 2. DOCSIS Transport (downstream)

Before delivery to the end user, IP packets are reconstructed by a DOCSIS modem (also known as a cable modem). The cable modem termination system (CMTS) at the headend, and the cable modem at the subscriber's end, make the cable network transparent to the user and make the user feel as if he/she is using a local area network (LAN). This also will enable multiplexing audio/video packets and IP packets in the same physical channel, if necessary.

In the MPEG-2 systems syntax and semantics are defined for multiplexing multiple programs in a single bitstream; the transport layer defines the semantics (or mechanism) for real-time delivery of multiple programs over error-prone channels. MPEG-2 Systems include the following functions:

- Timing and synchronization. The transmission of timing information in transport packets allows the receiver to synchronize with the encoder clock, which in turn helps audio and video to synchronize, avoiding lip-synch errors, etc.

- Packetization. The segmentation and encapsulation of elementary data streams into 188-byte transport packets. Included with each packet is the 4-byte packet header, which allows easy identification of each packet at the output of the transport demultiplexer.

- Multiplexing. The mechanisms to interspace transport packets of various elementary streams and program specific information (PSI) into a serial bitstream that complies with the MPEG-2 T-STD (transport stream decoder) model. This means timed delivery of the audio/video packets. The PSI information in the bitstream is used by the demultiplexer to demultiplex elementary streams uniquely at the decoder.

- Conditional Access. Provision for inclusion of access control information in the transport multiplex.

In the system/transport stream layer, bitstreams are split into 188-byte packets (a 184-byte payload and a 4-byte header) as shown in Figure 1. The header carries various information fields as shown in Figure 3. The PID (packet identifier), the most important field, has a length of 13 bits. The PID is a unique integer number associated with an elementary stream in a single or multi-program transport stream. The packets can carry various video and audio channels and other information, such as synchronization and timing, encryption, program information, access control, etc. The PID numbers help sort the packets into the specific streams to which they belong.

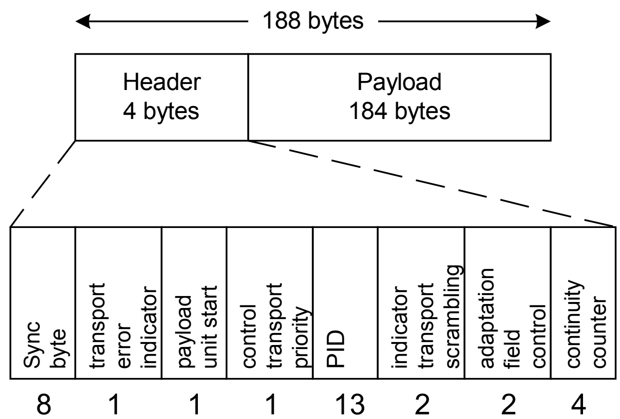
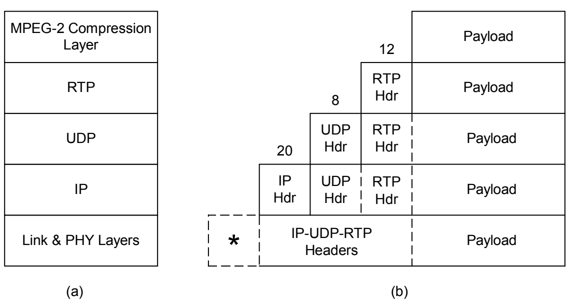


Figure 3. MPEG-2 Transport Packet

RTP—A Transport Protocol for Real-time Applications

The Moving Picture Expert Group (MPEG) has been active in creating international standards for compressed digital audio and video, and for the delivery mechanism (MPEG-2 transport)—a parallel effort was started by the Internet Engineering Task Force (IETF) for transport of data across various heterogeneous networks for non-real-time delivery. The transport standard created by IETF is known as Internet Protocol (IP). This protocol is widely used today for routing data across wide area networks (WAN), which may encompass several heterogeneous networks covering the entire world.

Various application-specific protocols were developed to take advantage of the IP protocol. TCP and UDP are the two dominant ones that sit on top of the IP layer. TCP is a connection-oriented protocol and, therefore, has additional, specific messages—a protocol for applications to request distant connections and a means for destinations to identify that they are ready to receive incoming connection requests. UDP provides a connectionless, unreliable service [10]. TCP was developed for guaranteed delivery of packets, whereas UDP does not. TCP is not well suited for real-time delivery of data because re-transmission conflicts with timed delivery of data. Timed delivery is an important requirement for real-time transport of audio and video. For that reason, UDP is the preferable mode of transmission for real-time applications. Below the IP and UDP headers, data-specific information is needed to convey payload-related (i.e., video, audio, etc) information. This additional information adds to the amount of the payload overhead. The protocol developed on top of actual payload of video, audio and data is known as real-time transport protocol (RTP) [4–6]. Figure 4a shows the generic packet format for RTP. Figure 4b provides an estimate of overhead from various layers such as IP-UDP-RTP. The bottom row of Figure 4b shows the payload and total overhead for a typical RTP packet.



*A media and protocol dependent header. For Ethernet the header size is 14 bytes.

Figure 4. RTP/UDP/IP Packet Layers

RTP may be used for unicast or multicast network services. Services provided by RTP include payload type identification, sequence numbering, time stamping and delivery monitoring. RTP does not provide a mechanism to ensure timely delivery or provide other quality-of-service guarantees, but relies on lower-layer services to do so. It does not guarantee delivery or prevent out-of-order delivery, nor does it assume that the underlying network is reliable and delivers packets in sequence.

RTP consists of two distinct, closely linked parts:

- The real-time transport protocol (RTP), to carry data that has real-time properties.
- The RTP control protocol (RTCP), to monitor quality-of-service and to convey information about the participant in an on-going session. Basically, it is a feedback mechanism from the receiver to the sender, related to the number of packets lost, etc. For example, loss of packets may be proportional to traffic congestion of a network segment; the server may take appropriate action to minimize packet loss.

The fixed header size of the RTP packet [4], [6] is 12 bytes, as shown in Figure 5. No header compression is assumed here. To reduce overhead, the size of the payload in a RTP packet can be increased, as range of payload size is 64 bytes to 1518 bytes. But if a packet with a large payload of compressed audio or video is lost, it may create respective audio/visual artifacts. Instead, video and audio of same time duration may be bundled in the same packet to keep the packet to an optimum size so that overhead is tolerable. This type of packaging is known as bundling or generic multiplexing of elementary streams in the same packet [9]. An extra four bytes in the header are required to indicate an offset of the audio payload and other audio-related information.

RTP Packet Header

The first eight bytes of the RTP header are:

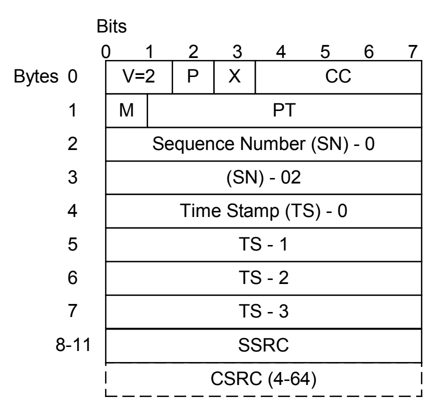


Figure 5. RTP Packet Header

The next four bytes comprise the synchronization source (SSRC) identifier and the next four to 64 bytes may comprise the contributing source (CSRC) identifiers.

RTP Payload

The payload size depends upon the audio-video codec type. For most standardized codecs, there are also payload headers [5], [8], as mentioned below, that immediately follow the fixed RTP header. As an example, the payload header for JPEG (RFC 2435, October 1998, RFC 2035, September 1996) streams consist of a 8-byte-long header, called the “main JPEG header,” followed by other related headers, such as, Restart Marker header, Quantization Table header, etc. Similarly there are payload headers for H.263+ (RFC 2429, October 1998), H.263 (RFC 2190, September 1997), H.261 (RFC 2032, October 1996), MPEG-1 / MPEG-2 Video /Audio (RFC 2250, January 1998), etc.

The RTP payload size depends upon the frame-sizes of the access units (compressed frames). RFC1889, January 1996 (RTP: A Transport Protocol for Real-Time Applications) defines the default packetizing interval for various audio streams as follows:

- Packetizing interval is 20 ms for frame-type codecs with a framing interval of 20 ms or less
- It shall be the framing interval, when the framing interval exceeds 20 ms. For G.723.1, the framing interval is 30 ms, so 30 ms shall be the packetizing interval.
- For G.729 codecs, the framing interval is 10 ms, so an RTP packet has two frames.
- For non-frame-type codecs, such as G.711 or G.726, the packetizing interval is 20 ms, etc.

UDP Header

The datagram checksum is two bytes long and validates the message contents [10], [6] as shown in Figure 6.

4	Source Port Number	Destination Port Number
4	Datagram Length	Datagram Checksum
8 bytes		

Figure 6. UDP Header

IP Datagram Header

The IP Header shown below consists of 20 bytes. There may be optionally another 0 to 40 bytes may be present in the header. Details can be found in [6] and [11].

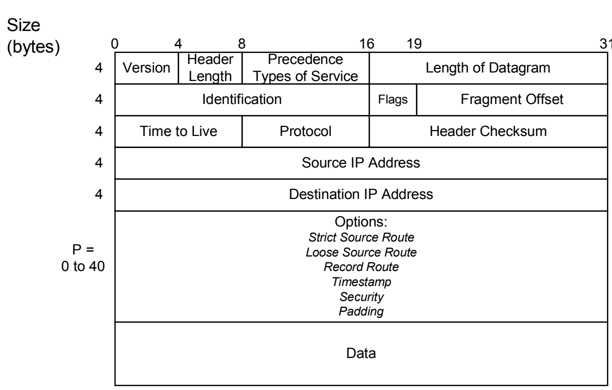


Figure 7. IP Datagram Header

Compression of Header Data

The method for compression of header data is known as payload header suppression (PHS). PHS is a process in which a portion of repeated MAC header information is suppressed. This provides a mechanism for avoiding redundant payload header data transmission. PHS is a form of data compression where efficiency of packet transmission increases; efficiency is gained by replacing a significant portion of the MAC packet header with few bytes.

A compression scheme [7] for the IP/UDP/RTP Headers (RFC 2508, February 1999) is initially targeted for applications sending audio and video over 14.4 and 28.8 dialup modems that provide full-duplex communication. This also may be used with reduced performance on simplex links. This compression scheme performs best on local links with low round-trip time.

The 12-byte long RTP packet header can be compressed to two to four bytes on an end-to-end basis. On a link-by-link basis, the combined IP+UDP+RTP header of 40 bytes can be compressed to two bytes for packets with no UDP Checksums, and to four bytes with Checksum. On a simplex link, or links with high delay, periodic refreshes with an uncompressed packet header are needed to restore the compression state in case of error. The link layer also must be able to provide an indication of switching between uncompressed and compressed header formats.

Almost half of the bytes in the IP and UDP headers may remain constant over the life of the connection. After sending the uncompressed header once, these fields may be removed from the compressed headers that follow. Differential coding on the changing fields in the remaining headers also reduces data size. For several fields that change in every packet, the difference from packet to packet is often constant, and its second-order difference is zero. When the uncompressed header and the first-order differences in the session state are shared between the compressor and decompressor, an indication of zero for second-order differences is sufficient for the decompressor to reconstruct the original header without any loss of information. This can be accomplished simply by adding first-order differences to the saved uncompressed header while each compressed packet is received.

While header suppression is an attractive technique to reduce overhead, it also has disadvantages. If a packet with an uncompressed header (reference packet) is lost, the follow-on packets with header suppression will be lost as well. To minimize reference packet loss, extra care needs to be taken, such as assigning a higher priority level or increasing robustness against bit errors in the header by adding FEC, etc.

IP Transport of Audio/Video over HFC Network

IP transport provides Internet connectivity for cable subscribers and is going to stay. Question arises if IP (RTP/UDP/IP) can be used as baseband transport for delivery of DTV signals. This way the entire HFC network will have one homogeneous transport and may provide some advantages such as provisioning, management, billing, etc. Before we make any such decision lets analyze such scenario. Figure 8 shows a typical block diagram of such an implementation. A packet framer packs MPEG-2 compressed audio/video in a RTP packet of desired size (e.g., L bytes). Actual payload will be L bytes minus the overhead (the headers). L bytes are broken into N MPEG-2 packets before input to the QAM modulator. For MPEG-2 encapsulation, the overhead is increased by N*4 bytes. No header compression is assumed here.

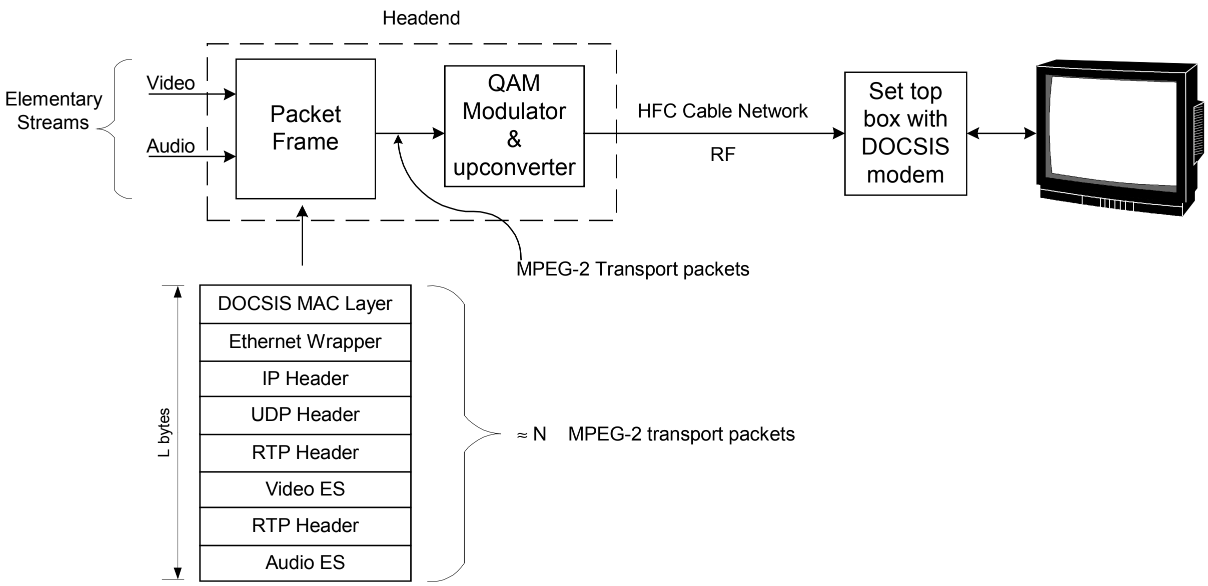


Figure 8. Typical scenerio for broadcast of DTV signal over DOCSIS channel (downstream)

Comparison of MPEG-2 and IP Transports

It may be observed from the above discussion that MPEG-2 transport is a link layer protocol. It has been created by MPEG for delivery of compressed digital audio/video in an isochronous (timely delivery of packet) mode. This protocol is particularly suited to broadcast delivery of multiple digital television programs in a very efficient way while providing synchronization between audio and video (no lipsync artifact). On the other hand, IP protocol has been created by IETF for end-to-end delivery of data in non-real-time over heterogeneous networks. IP is a higher layer protocol than MPEG-2 and addresses delivery of data on one-to-one or one-to-a-group basis, etc. It implies that IP packets are routable over WAN (wide area network) whereas MPEG-2 packets are not. For delivery of MPEG-2 compressed audio and video in real-time on the end-to-end basis IP protocol may be used to encapsulate audio and video data. To facilitate this, additional layers have been added over IP and the combined protocol is known as real time protocol (RTP). RTP is designed on top of UDP and IP, which is discussed before. Adding more layers provides flexibility but causes increase in overhead and efficiency of the transport protocol goes down. Efficiency may be defined as the ratio of actual payload divided by payload plus overhead. A brief comparison of two protocols is given in the Table 1.

The cable network spectrum is divided into a number of 6-MHz physical channels. When digitally modulated, a 6-MHz channel provides a digital bandwidth of 27 Mbps for 64 QAM and 38 Mbps for 256 QAM. For various bitrates of MPEG-2 compressed video, computations have been performed to find the number of logical channels that can be delivered over a 6-MHz physical channel using the two protocols.

		MPEG-2 Transport	RTP / UDP / IP Transport
1.	Ref. to OSI Layers	MPEG-2 is a link layer protocol. Link layer is the layer 2 of OSI.	RTP is a higher layer protocol. RTP is created by adding 3 layers above the link layer. IP is layer 3 of OSI and RTP resides over UDP/IP.
2.	Packet Size	Fixed 188 bytes.	Variable packet size (64 to 1518 bytes)
3.	Overhead	4 bytes for 184 byte payload.	40 bytes for 64 –1518 bytes. Overhead is much higher for comparable MPEG-2 packet size. Large packet size is avoided as packet loss can cause artifacts. Overhead may be reduced by bundling audio and video in the same packet
4.	Delivery	Optimized for broadcast delivery of compressed audio/video content. By adding MPEG-2 system layer conditional access mechanism delivery to one or a selected number of receivers is possible in an intranet like HFC cable network.	IP is optimized for one-to-one ore one-to-many non-real-time data delivery. RTP is an extension of IP for real-time use over Intranet. Not as efficient as MPEG-2 in a broadcast like application.
5.	Isochronous Delivery	MPEG-2 T-STD buffer model is designed for such a delivery.	No such buffer model is specified for RTP. Buffer model depends on application.
6.	Audio/Video Synchronization	Keeps tight synchronization between Audio and Video. Proven through large-scale implementation in broadcast applications.	A few issues to be solved for synchronized delivery of A/V over internet. A QOS capable network may be of some help.
7.	Routability	Not designed for routing over heterogeneous networks.	Designed for routing over internet.
8.	Efficiency	97.7%	< 85%

Table 1. Comparison of MPEG-2 and IP Protocols

Table 2 presents the computational result for a 64-QAM modulated channel while Table 3, for 256 QAM no header compression is assumed in putting the number of channels for IP. It may be observed from Tables 2 & 3, that MPEG-2 transport is preferable to IP as the former provides more channels. More channels mean more revenue. The next question comes about streaming video and audio. Streaming from the headend (such as VOD) to any cable subscriber may be implemented using either MPEG-2 transport or IP transport. MPEG-2 transport will be preferable over IP from the reason of overhead. When a cable subscriber wants to reach a source located outside the cable network for a streaming video, IP (RTP/UDP/IP) will prevail as packets constituting the stream need routing over the internet. Internet is an unpredictable network where loss of packet or variable packet delay may occur. Near isochronous delivery may be achieved using some form of QOS or a private backbone.

Encoding Resolution	Bit Rate (Mbps)	MPEG-2 Transport Logical Channels	Video Over IP Logical Channels
352x240 (CIF) (Movies)	1.5	17	14
352x480 (Half) (Movies)	2	13	11
540x480 (3/4) (Movies)	3	8	7
704x480 (Full) (Movies)	4	6	5
540x480 (3/4) (Sports)	5	5	4
704x480 (Full) (Sports)	6	4	3
HDTV 1080x1920 (Full)	19	1	1

Table 2. Logical channels for MPEG-2 Transport and Video over IP (64 QAM 27 Mbps)

Encoding Resolution	Bit Rate (Mbps)	MPEG-2 Transport Logical Channels	Video Over IP Logical Channels
352x240 (CIF) (Movies)	1.5	25	21
352x480 (Half) (Movies)	2	19	15
540x480 (3/4) (Movies)	3	12	10
704x480 (Full) (Movies)	4	9	7
540x480 (3/4) (Sports)	5	7	6
704x480 (Full) (Sports)	6	6	5
HDTV 1080x1920 (Full)	19	2	1

Table 3. Logical channels for MPEG-2 Transport and Video over IP (256 QAM 38 Mbps)

The MPEG-4 Standard and the Related Content

MPEG-4 is a compression standard and no specific transport protocol has been created to deliver MPEG-4 compressed elementary streams. MPEG-2 Systems standard has been amended to carry MPEG-4 content both at elementary stream level and as multiplexed one. Also MPEG and IETF (Audio Video Transport group) have been working together to create a protocol to deliver MPEG-4 content over RTP (RTP/UDP/IP). In this regard a few RFC have been submitted for consideration.

Conclusion

Baseband transport of digital content over the HFC cable network has been analyzed. The two leading transport protocols, MPEG-2 and IP, have been studied. MPEG-2 transport is very efficient for broadcast delivery of MPEG-2 compressed audio/video and private data. IP has a larger overhead, but provides connectivity to the internet, a world outside of the cable intra-network. It is expected that MPEG content delivery and internet connectivity will be two important parts of the cable business. Broadcast of MPEG content from the headend is very efficient using MPEG-2 transport, while providing internet connectivity via HFC cable network has become a necessity. It may be concluded that both transport protocols are going to coexist in the HFC cable network for a while before one takes over the other.

References

[1] ISO/IEC IS 13818-1, “Information Technology—Generic coding of Moving Pictures and Associated Audio Information: Systems.”

[2] ISO/IEC 11172-1, “Coding of Moving Pictures and Associated Audio for Storage Media at up to about 1.5 Mbit/sec.” Part I; Systems.

[3] Annex B to ITU-T Recommendation J.83 (10/95), Digital multi-programme systems for television sound and data services for cable distribution.

[4] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” RFC 1889, January 1996.

[5] M. Reha Civanlar, Glenn L. Cash, Barry G. Haskell, “RTP Payload Format for Bundled MPEG,” draft-civanlar-bmpeg-00.txt, August 1996.

[6] Douglas E. Comer, “Computer Networks and Internets,” Prentice Hall, ISBN 0-13-083617-6.

[7] S. Casner, V. Jacobson, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links," draft-ietf-avt-crtp-05.txt, July 27 1998.

[8] M. Handley, "Guidelines for writers of RTP payload format specifications," Internet Draft.

[9] M. Handley, "GeRM: Generic RTP Multiplexing," draft-ietf-avt-germ-00.txt, November 11, 1998.

[10] J. Postel, "User Datagram Protocol," RFC 0768, August 1980.

[11] Defense Advanced Research Projects Agency (DARPA), "INTERNET PROTOCOL, DARPA INTERNET PROGRAM, PROTOCOL SPECIFICATION", RFC 0791, September 1981.

[12] Data Over Cable Service Interface Specification, Radio Frequency Interface Specification, DOCSIS 1.0 and DOCSIS 1.1.

Acknowledgements

The authors would like to thank the management of CableLabs and Philips Semiconductor for their support in performing this research. Special thanks are due to Mr. Rouzbeh Yassini of YAS Corporation and other members of the team.